Available online at www.sciencedirect.com
ScienceDirect

Journal of Approximation Theory 156 (2009) 82–94

Journal of Approximation Theory

www.elsevier.com/locate/jat

Perturbations of Jordan matrices

E.B. Davies [a,*], Mildred Hager [b,c]

[a] Department of Mathematics, King's College, Strand, London WC2R 2LS, United Kingdom
[b] Department of Mathematics, University of California, Berkeley, CA 94720, United States
[c] Matematikcentrum, Universitet Lund, Box 118, 22100 Lund, Sweden

Received 17 July 2007; received in revised form 18 March 2008; accepted 5 April 2008
Available online 20 November 2008

Communicated by Serguei Denissov

Abstract

We consider perturbations of a large Jordan matrix, either random and small in norm or of small rank. In the case of random perturbations we obtain explicit estimates which show that as the size of the matrix increases, most of the eigenvalues of the perturbed matrix converge to a certain circle with centre at the origin. In the case of finite rank perturbations we completely determine the spectral asymptotics as the size of the matrix increases.

Keywords: Jordan matrices; Eigenvalues; Perturbations; Lidskii; Random matrices; Spectrum; Pseudospectrum

1. Introduction

The analysis of the eigenvalues of highly non-self-adjoint matrices and operators is now a well developed field. Such operators appear frequently in the study of certain PDEs [1] (e.g. in fluid dynamics), as well as their numerical discretisations, and the understanding of their spectral behaviour is fundamental for stability studies [2,3]. A particularly challenging aspect is the perturbation theory, because perturbations can have a much bigger effect on the eigenvalues of such operators and matrices than in the self-adjoint case [4].

From a theoretical point of view, some general facts are known. If $A = B + \delta K$, then any non-degenerate eigenvalue of B yields an eigenvalue of A that depends analytically on δ up to the point at which two eigenvalues cross [5]. In the non-self-adjoint case B might have degenerate

* Corresponding author.
E-mail addresses: E.Brian.Davies@kcl.ac.uk (E.B. Davies), hager@maths.lth.se (M. Hager).

doi:10.1016/j.jat.2008.04.021

eigenvalues which are associated with groups of eigenvalues of A whose description can be much more complicated. An important article of Lidskii [6,7] explained in detail what may happen for small enough values of δ. Generically one obtains a Puiseux series for the eigenvalues if the perturbation parameter is small enough. However, for small enough but non-generic Hessenberg perturbations the eigenvalues may split into several rings, [8].

In this paper we go beyond Lidskii's theory in two different directions, whose methodologies are related. We assume throughout that the unperturbed matrix is an $N \times N$ Jordan block. This simple special case has been investigated in considerable detail [3,8], because it is the starting point for the understanding of the behaviour of degenerate eigenvalues of more general matrices under small perturbations. In the present work, we shall consider larger perturbations, and focus on the asymptotic behaviour as $N \to \infty$ (although our main theorems are valid for all $N \geq 2$). In the first part of the paper we assume that K is a generic, i.e. random, perturbation and that the perturbation parameter is small, but too large for the applicability of Lidskii's theory. In the second part of the paper we study a perturbation of small rank but with norm of order 1.

Although Lidskii's theory is not applicable, it is very surprising that numerical calculations indicate that in both cases most of the eigenvalues are still close to the Lidskii circle, but that a small proportion are far away from it. We prove that most of the eigenvalues lie close to the Lidskii circle in the random and finite rank cases. In the random case we obtain a probabilistic upper bound on the number of exceptional eigenvalues, all of which lie inside the Lidskii circle, but do not obtain any information about their asymptotic distribution; of course, since the problem is probabilistic, the actual number of such eigenvalues will vary from sample to sample, and for some such perturbations there will be none. In the finite rank case we give an exact procedure for computing the exceptional eigenvalues, which may lie inside or outside the Lidskii circle. In both cases our theorems are asymptotic in character, and we do not claim that the estimates obtained are numerically sharp.

A quantitative measure of spectral instability is provided by the notion of pseudospectra, which becomes interesting when the matrix involved is far from being normal; see [9,10] for detailed discussions and many references. If $\delta > 0$ the δ-pseudospectra of an operator B are defined by

$$\begin{aligned} \mathrm{Spec}_\delta(B) &:= \mathrm{Spec}(B) \cup \{z \notin \mathrm{Spec}(B) : \|(B - zI)^{-1}\| > \delta^{-1}\} \\ &= \bigcup_{\{K:\|K\|<1\}} \mathrm{Spec}(B + \delta K)\,, \end{aligned} \tag{1}$$

where Spec denotes the spectrum of a matrix. For a non-self-adjoint matrix B, the norm of the resolvent $(B - zI)^{-1}$ can be much larger than in the self-adjoint case. Hence the perturbed eigenvalues can be quite far away from the unperturbed ones: the second equality in (1) implies that a perturbation of B of size δ can move the eigenvalues anywhere inside $\mathrm{Spec}_\delta(B)$. In particular, computed eigenvalues of a large matrix may be very inaccurate if $\mathrm{Spec}_\delta(B)$ is a large region, where δ is the rounding error of the computations. In this note, we study this phenomenon in some detail for the Jordan block matrix, perturbed either by a matrix of small rank, in which case the analysis is much sharper, or by a random matrix with a small norm. The problem studied in this paper was proposed by Zworski, who showed one of us how the general methods of Sjöstrand and Zworski in [11] could be adapted to this particular setting.

We define the standard $N \times N$ Jordan matrix J by

$$(J)_{r,s} := \begin{cases} 1 & \text{if } s = r + 1, \\ 0 & \text{otherwise,} \end{cases} \tag{2}$$

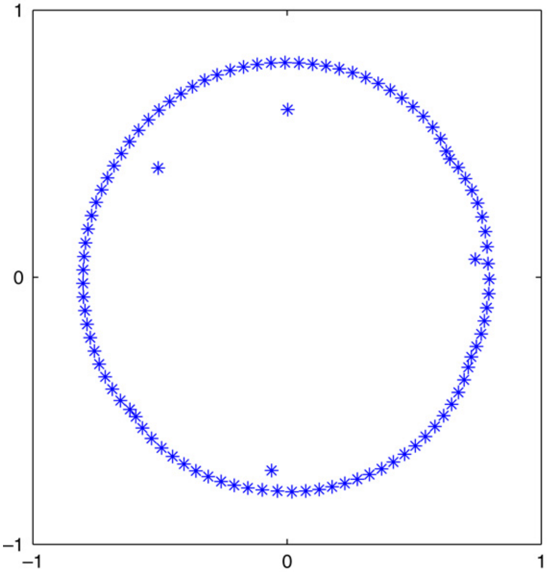

Fig. 1. Eigenvalues of a small random perturbation of the Jordan matrix J, with $N = 100$ and $\delta = 10^{-10}$.

where $r, s = 1, \ldots, N$. Fig. 1 shows the results of a MATLAB computation of $\mathrm{Spec}(J + \delta K)$, when $N = 100$, $\delta = 10^{-10}$ and K is a complex Gaussian random matrix. Most of the eigenvalues accumulate around a circle (the Lidskii circle) with centre at the origin; however four eigenvalues lie well inside the circle. This cannot happen in Lidskii's theory, but the latter is not applicable in this context. In Section 2, we prove that most of the eigenvalues lie close to the Lidskii circle with high probability in the asymptotic regime $N \to \infty$. The section illustrates the methods and ideas of [12] in a very concrete setting. It is worth noting that the results in this section only apply for small values of δ. In numerical experiments one finds that for fixed but reasonably large N and generic random K the eigenvalues of $J + \delta K$ are randomly distributed if $\delta > 1$.

If one adds a strictly upper triangular matrix to J then the spectrum is not changed. In Section 3, we concentrate on large, non-random perturbations whose non-zero entries are all close to the bottom-left-hand corner of the matrix. We give a complete asymptotic analysis as $N \to \infty$ of the spectrum for all such perturbations. The equation to be solved is written down in Theorem 8. The asymptotic forms of the solutions to this equation are described in Theorems 9 and 11.

Before continuing we mention two further papers, [13,14], that deal with other types of perturbations of Jordan blocks.

2. Small random perturbations of Jordan matrices

Throughout this paper $e_1, e_2, \ldots, e_N$ denotes the standard basis of column vectors in $\mathbf{C}^N$. The norm of an $N \times N$ matrix A is defined by

$$\|A\| := \max\{\|Av\| : v \in \mathbf{C}^N \text{ and } \|v\| \leq 1\},$$

where $\|v\|$ is the Euclidean norm of v. Since the quantity $N^{c/N}$ occurs repeatedly below, with various values of c, it is worth commenting that it converges to 1 as $N \to \infty$. The non-random parameter $\delta > 0$ at the centre of our arguments is understood to depend on N and to satisfy certain inequalities that are specified as needed. We also put $R := \delta^{1/N}$; as a special case one may choose an N-independent constant $R \in (0, 1)$ and put $\delta := R^N$. We use the notation $D(r) := \{z \in \mathbf{C} : |z| \le r\}$ and refer to $\{z : |z| = R\}$ as the Lidskii circle.

Although we are primarily interested in the spectral properties of $J + \delta K$ in the asymptotic limit $N \to \infty$, our bounds are valid for all $N \ge 2$. However, we have not attempted to obtain the optimal constants in the bounds. Our main theorems below state, in a quantitative form, that for large N most of the eigenvalues of $J + \delta K$ are close to the Lidskii circle, but a small proportion may be well inside it.

Theorem 1. *Let $\alpha,\ \beta \ge 0$ and put $\gamma := 2 + 2\alpha + \beta$. Given $N \ge 2$, define J by (2) and let K be an $N \times N$ matrix satisfying $\|K\| \le N^\alpha$ and $|K_{N,1}| \ge N^{-\beta}$. Then if $0 < \delta < \frac{1}{4} N^{-\gamma}$, $R := \delta^{1/N}$ and $\sigma > 0$, we have*

$$\operatorname{Spec}(J + \delta K) \subseteq D(R\, N^{(1+\alpha)/N}) \tag{3}$$

and

$$\#\left(\operatorname{Spec}(J + \delta K) \cap D(R\mathrm{e}^{-\sigma})\right) \le \frac{2}{\sigma} + \frac{1 + \alpha + \beta}{\sigma} \ln(N). \tag{4}$$

The hypotheses on K are satisfied with high probability for a wide variety of randomly generated matrices. The following theorem is a typical case.

Theorem 2. *Given $N \ge 2$, let K be an $N \times N$ random matrix with entries independently and identically distributed according to a complex Gaussian law centred at 0 and of variance 1. Then if $0 < \delta \le \frac{1}{4} N^{-7}$, $R := \delta^{1/N}$ and $\sigma > 0$, with probability at least $1 - 2N^{-2}$ one has*

$$\operatorname{Spec}(J + \delta K) \subseteq D(R\, N^{3/N}) \tag{5}$$

and

$$\#\left(\operatorname{Spec}(J + \delta K) \cap D((R\mathrm{e}^{-\sigma}))\right) \le \frac{2}{\sigma} + \frac{4}{\sigma} \ln(N)\ . \tag{6}$$

This gives an explicit upper bound on the number of eigenvalues in discs of varying radii.

The remainder of the section is devoted to the proofs of these theorems.

Lemma 3. *If $\alpha \ge 0$, $\|K\| \le N^\alpha$ and $0 < \delta \le \frac{1}{4} N^{-1-\alpha}$ then*

$$\operatorname{Spec}(J + \delta K) \subseteq D(RN^{(1+\alpha)/N}).$$

Proof. The formula

$$(J - zI)^{-1}_{r,s} = \begin{cases} -z^{r-s-1} & \text{if } 1 \le r \le s \le N, \\ 0 & \text{otherwise,} \end{cases}$$

valid for all $z \ne 0$, implies that

$$\|(J - zI)^{-1}\| \le |z|^{-1} + |z|^{-2} + \cdots + |z|^{-N} \le \begin{cases} N|z|^{-1} & \text{if } |z| \ge 1, \\ N|z|^{-N} & \text{if } |z| \le 1. \end{cases}$$

If $|z| \geq 1$ then

$$\|\delta K(J - zI)^{-1}\| \leq \frac{1}{4} N^{-1-\alpha} N^{\alpha} N < 1,$$

so $z \notin \mathrm{Spec}(J + \delta K)$ by a standard argument using the Neumann series. If $RN^{(1+\alpha)/N} < |z| < 1$ then

$$\|\delta K(J - zI)^{-1}\| \leq \delta N^{\alpha} N |z|^{-N} < 1,$$

so once again $z \notin \mathrm{Spec}(J + \delta K)$. □

We next apply Schur's complementation formula to a special setting called a Grushin problem. These problems are widely used in spectral theory to relate the eigenvalues of an operator to the zeros of some holomorphic function (see [11, Sect 2.2]).

Lemma 4. *Let A be an $N \times N$ matrix and define the $(N + 1) \times (N + 1)$ block matrix $\mathcal{J}$ by*

$$\mathcal{J} := \begin{pmatrix} A & e_N \\ e_1' & 0 \end{pmatrix}, \tag{7}$$

where e_1' is the transpose of e_1. Then $\mathcal{J}$ is invertible if and only if the matrix $\tilde{A}$ obtained by deleting the left-hand column and bottom row of A is invertible. Assuming that this is the case put

$$\mathcal{E} := \mathcal{J}^{-1} := \begin{pmatrix} E & F \\ G & H \end{pmatrix}, \tag{8}$$

where E is an $N \times N$ matrix, F is an $N \times 1$ matrix, G is a $1 \times N$ matrix and $H \in \mathbf{C}$. Then A is invertible if and only if H is non-zero.

Proof. Direct calculations show that

$$\det(\mathcal{J}) = (-1)^N \det(\tilde{A})$$

and

$$H = (-1)^N \frac{\det(A)}{\det(\tilde{A})}.$$

Explicitly, if $H \neq 0$, $A^{-1} = E - FH^{-1}G$, which is Schur's complementation formula. □

From this point onwards we assume that $A = J + \delta K - zI$ and attach the subscripts δ, z to $\mathcal{J}$, $\mathcal{E}$, E, F, G and H. Thus the zeros of $H_{\delta,z}$ as a function of z determine the spectrum of $J + \delta K$.

Example 5. If $\delta = 0$ then $\det(\tilde{A}) = 1$, so $\mathcal{J}_{0,z}$ is invertible for every $z \in \mathbf{C}$. Direct calculations yield

$$\begin{aligned} \left(E_{0,z}\right)_{r,s} &= \begin{cases} z^{r-s-1} & \text{if } s < r, \\ 0 & \text{otherwise,} \end{cases} \\ \left(F_{0,z}\right)_{r,1} &= z^{r-1}, \\ \left(G_{0,z}\right)_{1,s} &= z^{N-s}, \\ H_{0,z} &= z^N \end{aligned}$$

for all relevant r, s. Assuming $|z| \le 1$ we deduce that

$$\begin{aligned} &\|\mathcal{E}_{0,z}\| \le N+1, \qquad \|E_{0,z}\| \le N, \qquad \|F_{0,z}(z)\| \le N^{1/2}, \qquad \|G_{0,z}(z)\| \le N^{1/2}, \\ &\|\mathcal{E}_{0,0}\| = \|E_{0,0}\| = \|F_{0,0}\| = \|G_{0,0}\| = 1, \qquad H_{0,0} = 0. \end{aligned} \tag{9}$$

Our next task is to obtain some bounds on $H_{\delta,z}$ that will provide information about the location of its zeros.

Lemma 6. *Let $\alpha, \beta \ge 0$, $\gamma := 2+2\alpha+\beta$, $\|K\| \le N^\alpha$, $|K_{N,1}| \ge N^{-\beta}$, $N \ge 2$, $0 < \delta < \frac{1}{4}N^{-\gamma}$ and $|z| \le R := \delta^{1/N}$. Let $H_{\delta,z}$ be the complex number H of (8) if one puts $A := J + \delta K - zI$. Then*

$$|H_{\delta,z}| \le 3R^N N^{1+\alpha} \tag{10}$$

and

$$|H_{\delta,0}| \ge \frac{1}{2} R^N N^{-\beta}. \tag{11}$$

Proof. We first observe that $\mathcal{J}_{\delta,z} = \mathcal{J}_{0,z} + \delta\mathcal{K}$ where

$$\mathcal{K} := \begin{pmatrix} K & 0 \\ 0 & 0 \end{pmatrix}.$$

The Neumann expansion for $\mathcal{E}_{\delta,z} = \mathcal{J}_{\delta,z}^{-1}$, namely

$$\mathcal{E}_{\delta,z} = \mathcal{E}_{0,z} - \mathcal{E}_{0,z}\delta\mathcal{K}\mathcal{E}_{0,z} + \mathcal{E}_{0,z}\delta\mathcal{K}\mathcal{E}_{0,z}\delta\mathcal{K}\mathcal{E}_{0,z} - \cdots,$$

is norm convergent under the stated conditions on δ and z because

$$\|\delta\mathcal{K}\mathcal{E}_{0,z}\| \le \delta\|K\|(N+1) \le \delta\|K\|2N < 1$$

by combining the assumptions of the lemma with (9). This implies in particular that the series

$$H_{\delta,z} = H_{0,z} - G_{0,z}\delta K F_{0,z} + G_{0,z}\delta K E_{0,z}\delta K F_{0,z} - \cdots \tag{12}$$

is also convergent.

If $|z| \le R$ then by combining the assumptions of the lemma with (9), we obtain

$$\begin{aligned} |H_{\delta,z}| &\le R^N + N\delta\|K\| + (N\delta\|K\|)^2 + (N\delta\|K\|)^3 + \cdots \\ &\le R^N + 2N\delta\|K\| \\ &\le 3R^N N^{1+\alpha}. \end{aligned}$$

Expansion (12) with $z = 0$ also implies that

$$\begin{aligned} |H_{\delta,0}| &\ge \delta N^{-\beta} - (N\delta\|K\|)^2 - (N\delta\|K\|)^3 - \cdots \\ &\ge \delta N^{-\beta} - 2(N\delta\|K\|)^2 \\ &\ge \delta\left(N^{-\beta} - 2N^{2+2\alpha}\delta\right) \\ &\ge \frac{1}{2} R^N N^{-\beta}. \quad \square \end{aligned}$$

To estimate the number of zeros of the holomorphic function $f(z) := H_{\delta,z}$, we will need the next proposition.

Proposition 7 (*The Poisson–Jensen Formula*). *Let f be a bounded holomorphic function on $D(R)$, where $0 < R < \infty$. Let M be the number of zeros of f in $D(R\mathrm{e}^{-\sigma})$ for some positive constant σ. Then*

$$M \leq \frac{1}{\sigma} \ln\left(\frac{\|f\|_{L^\infty(D(R))}}{|f(0)|}\right). \tag{13}$$

Proof. We assume that f does not vanish on $|z| = R$, since otherwise we may diminish R slightly and obtain our result by taking a limit over increasing radii fulfilling this assumption. The proposition is a direct consequence of formula $(1.2')$, p.163 in [15]: if f is a holomorphic function in $D(R)$ with zeros a_μ there, then

$$\ln|f(0)| = \frac{1}{2\pi}\int_0^{2\pi} \ln|f(R\mathrm{e}^{\mathrm{i}\theta})|\mathrm{d}\theta - \sum_{|a_\mu|<R} \ln\frac{R}{|a_\mu|} \leq \ln\|f\|_\infty - \sum_{|a_\mu|<R} \ln\frac{R}{|a_\mu|}. \tag{14}$$

Hence

$$-\ln|f(0)| + \ln\|f\|_\infty \geq \sum_{|a_\mu|<R} \ln\frac{R}{|a_\mu|} \geq \sum_{|a_\mu|<R\mathrm{e}^{-\sigma}} \ln\frac{R}{R\mathrm{e}^{-\sigma}} = M\sigma, \tag{15}$$

which is our proposition. $\square$

Proof of Theorem 1. Inclusion (3) was proved in Lemma 3. The invertibility of the matrix $\mathcal{J}_{\delta,z}$ for all $z \in D(R)$ was shown in the proof of Lemma 6, so Lemma 4 implies that the eigenvalues of $J + \delta K$ coincide with the zeros of $H_{\delta,z}$ in $D(R)$. The number of zeros of $H_{\delta,z}$ in $D(R\mathrm{e}^{-\sigma})$ is controlled by Proposition 7 and Lemma 6. These yield

$$\begin{aligned} M &\leq \frac{1}{\sigma}\ln\left(\frac{3R^N N^{1+\alpha}}{\frac{1}{2}R^N N^{-\beta}}\right) \\ &= \frac{1}{\sigma}\ln\left(6N^{1+\alpha+\beta}\right) \\ &\leq \frac{2}{\sigma} + \frac{1+\alpha+\beta}{\sigma}\ln(N). \quad \square \end{aligned}$$

Proof of Theorem 2. This depends on showing that a complex Gaussian random perturbation fulfills the assumptions of Theorem 1 with high probability.

If K is a random matrix with independent complex Gaussian normal distributed entries and $a > 0$ then

$$\begin{aligned} P[\|K\| > a] &\leq P\left[\sum_{j,k=1}^N |K_{jk}|^2 > a^2\right] \\ &\leq E\left[a^{-2}\sum_{j,k=1}^N |K_{jk}|^2\right] = N^2/a^2\,. \end{aligned} \tag{16}$$

Hence

$$P[\|K\| > N^2] \leq N^{-2}. \tag{17}$$

If $s > 0$,

$$P[|K_{N,1}| < s] = 1 - \exp\left(-\frac{s^2}{2}\right) \leq s^2 . \tag{18}$$

Therefore

$$P[|K_{N,1}| < N^{-1}] \leq N^{-2}.$$

Combining these two estimates proves that with probability at least $1 - 2N^{-2}$ we have both $\|K\| \leq N^2$ and $|K_{N,1}| \geq N^{-1}$. We may now apply Theorem 1 with $\alpha = 2$ and $\beta = 1$. $\square$

Note that bounds similar to (16) and (18) can be proved for a wide variety of other probability laws governing the coefficients of K.

3. Perturbations of boundary condition type

Let C be a $k \times k$ matrix and let $\delta > 0$ be the perturbation parameter. We consider the spectrum of

$$A := J + \delta K \tag{19}$$

where J is given by (2) and the $N \times N$ matrix K has the block form

$$K := \begin{pmatrix} 0 & 0 \\ C & 0 \end{pmatrix},$$

the (zero) top-right-hand entry being of size $(N - k) \times (N - k)$. We also assume that $N > 2k$.

From this point onwards we assume that $\delta := R^N$ for some $R \in (0, \infty)$. Surprisingly the asymptotic behaviour of $\mathrm{Spec}(A)$ as $N \to \infty$ keeping k and C fixed does not depend on the choice of R. If $0 < R < 1$ then δK is a very small perturbation of J, but for $R > 1$ the reverse holds. In contrast with the results in the last section for random perturbations, the same analysis applies in both cases.

The form of the perturbation looks artificial from the point of view of matrix analysis, but not if one thinks of A as generating discrete time dynamics on $\{1, \ldots, N\}$ in which the particle moves steadily from 1 to N. The matrix C then describes a general re-entry law, or boundary condition, which takes the particle back to the starting point. In a later paper one of us considers a similar problem in which the ordered interval $\{1, \ldots, N\}$ with re-entry law is replaced by a general directed graph; the geometry of the graph is then a vital ingredient when determining the structure of the spectrum [16]. Our results may have also connections with the analysis of para-orthogonal polynomials on the unit circle in [17].

Theorem 8. *Let A be defined by (19) where $\delta := R^N$ and $R \in (0, \infty)$. Then*

$$\mathrm{Spec}(A) = \{Rz : z^N = f(z)\}$$

provided $N > 2k$, where f is the (N-independent) polynomial

$$f(z) := \sum_{i,j=1}^{k} C_{i,j} (Rz)^{j-i+k-1}.$$

Proof. Let $g_r(\lambda)$ be the determinant of the $r \times r$ matrix M_r obtained from $\lambda I - J - \delta K$ by deleting its top $N-r$ rows and the leftmost $N-r$ columns. By expanding $\det(M_N)$ down the leftmost column we obtain

$$g_N(\lambda) = \lambda g_{N-1}(\lambda) - \delta C_{k,1} - \delta C_{k-1,1}\lambda - \delta C_{k-2,1}\lambda^2 - \cdots - \delta C_{1,1}\lambda^{k-1}.$$

There are similar formulae for $g_r(\lambda)$ if $N-k < r < N$, and one also has $g_{N-k}(\lambda) = \lambda^{N-k}$. The formula

$$\begin{aligned}\det(\lambda I - J - \delta K) &= g_N(\lambda)\\ &= \lambda^N - \delta \sum_{i,j=1}^{k} C_{i,j}\lambda^{j-i+k-1}\end{aligned}$$

follows inductively. The proof is completed by making the change of variables $\lambda := Rz$. □

4. The equation $z^N = f(z)$

Theorem 8 reduces the asymptotic analysis of the eigenvalues of A to the study of the equation $z^N = f(z)$, where f is a certain polynomial. From this point onwards we abandon the matrix problem and assume that f is a general analytic function, because the proofs carry through with only trivial alterations in this case. Since we only have to consider one such function, instead of a random ensemble of such functions, we are able to obtain a much more complete analysis.

The standing assumptions in this section are as follows. Let U be a region in the complex plane that contains $D(1+\delta)$ for some $\delta > 0$. Let f be a bounded analytic function defined on U. We assume that $f(z) = 0$ has h distinct solutions z_i satisfying $|z_i| < 1$, each with multiplicity m_i. We put

$$n := \sum_{i=1}^{h} m_i. \tag{20}$$

By reducing $\delta > 0$ we may assume that $|z_i| < 1-\delta$ for all i. We wish to determine the asymptotic distribution of the solutions of $z^N = f(z)$ as $N \to \infty$. Our first theorem provides the precise asymptotic location of the n exceptional solutions.

Theorem 9. *For every $\varepsilon \in (0, \delta)$ there exists N_ε such that if $N \geq N_\varepsilon$ then $z^N = f(z)$ has m_i solutions in the ε-neighbourhood of z_i for each $i \in \{1, \ldots, h\}$, no other solutions in $D(1-\varepsilon)$, no solutions in $U \setminus D(1+\varepsilon)$ and $N-n$ solutions in the annulus $\{z : 1-\varepsilon < |z| < 1+\varepsilon\}$.*

Proof. If N is large enough then $(1+\varepsilon)^N > \max\{|f(z)| : z \in U\}$, so the equation has no solutions in $U \setminus D(1+\varepsilon)$. By applying Rouche's theorem to $z^N - f(z)$ regarded as a small perturbation of z^N, we see that for all large enough N the equation $z^N = f(z)$ has N solutions inside $D(1+\varepsilon)$. A similar argument but regarding $f(z) - z^N$ as a small perturbation of $f(z)$, implies that the equation $z^N = f(z)$ has n solutions inside $D(1-\varepsilon)$, provided N is large enough, and that these converge to the zeros of $f(z)$ as $N \to \infty$. The remaining $N-n$ solutions must lie in the stated annulus. □

Example 10. We revert temporarily to the context and notation of Theorem 8. If $\delta := 1$, $k := 2$ and $C := \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then the eigenvalues of $J + K$ are the solutions of

$$z^N = c + (a+d)z + bz^2.$$

For large N most of these are close to the unit circle, but there are also isolated eigenvalues close to those solutions z of $c + (a + d)z + bz^2 = 0$ that satisfy $|z| < 1$. There may be 0, 1 or 2 isolated eigenvalues. If a, b, c, d are chosen randomly, one might be able to compute the expected number of isolated eigenvalues.

Returning to the context of Theorem 9, we wish to determine the asymptotic behaviour as $N \to \infty$ of the $N - n$ solutions of $z^N = f(z)$ in the annulus $\{z : 1 - \varepsilon < |z| < 1 + \varepsilon\}$, assuming that $f(z) \neq 0$ whenever $|z| = 1$. We put

$$f(\mathrm{e}^{\mathrm{i}s}) := \rho(s)\mathrm{e}^{\mathrm{i}\phi(s)}$$

where $\rho(s)$ is positive and periodic on $[0, 2\pi]$ while

$$\phi(2\pi) = \phi(0) + 2\pi n, \tag{21}$$

where n is defined in (20). Both $\rho(s)$ and $\phi(s)$ are real analytic functions of s. It is easy to see that if $N > \max\{\phi'(s) : s \in [0, 2\pi]\}$ the equation

$$\phi(s) = Ns \mod (2\pi) \tag{22}$$

has $N - n$ distinct solutions s_j in $[0, 2\pi)$. If these are labelled in increasing order then

$$s_{j+1} - s_j = 2\pi/N + O(1/N^2) \tag{23}$$

for all j. We will show that for all large enough N the solutions of $z^N = f(z)$ are very close to the points

$$a_j := \rho(s_j)^{1/N}\mathrm{e}^{\mathrm{i}s_j}. \tag{24}$$

We conjecture that the following theorem remains valid for $\alpha = 2$.

Theorem 11. *Define n by (20) and a_j by (24). If $\alpha \in (1, 2)$ and $f(z) \neq 0$ whenever $|z| = 1$, then there exists a constant $\overline{N}$ such that for all $N \geq \overline{N}$ and every $j \in \{1, \ldots, N-n\}$ the equation $z^N = f(z)$ has a solution z_j satisfying*

$$|z_j - a_j| \leq 3N^{-\alpha}.$$

To leading order the $N - n$ solutions of $z^N = f(z)$ that are close to the unit circle are uniformly distributed around it.

Proof. We focus on a particular value of j and leave the reader to verify that the bounds obtained are uniform with respect to j. If we put $w := \mathrm{e}^{-\mathrm{i}s_j}z$ then an elementary calculation using (22) shows that finding the solution of $z^N = f(z)$ closest to $\mathrm{e}^{\mathrm{i}s_j}$ is equivalent to finding the solution of $w^N = g(w)$ closest to 1, where

$$g(w) := \mathrm{e}^{-\mathrm{i}\phi(s_j)} f(\mathrm{e}^{\mathrm{i}s_j} w).$$

We define the sequence $u_m := r_m\mathrm{e}^{\mathrm{i}\theta_m}$ for $m = 1, 2, \ldots$ by

$$u_1 := 1, \quad u_{m+1} := \{g(u_m)\}^{1/N}, \tag{25}$$

where we always take the Nth root with the smallest argument. We will prove that this sequence converges to a solution of $u^N = g(u)$. Since

$$g(1) = \mathrm{e}^{-\mathrm{i}\phi(s_j)} f(\mathrm{e}^{\mathrm{i}s_j}) = \rho(s_j) > 0,$$

we see that $r_1 = 1$, $\theta_1 = \theta_2 = 0$ and

$$r_2 = \rho(s_j)^{1/N} = 1 + O(1/N). \tag{26}$$

In the following arguments c_j denote positive constants that do not depend on N. We will prove that if

$$S_N := \{r\mathrm{e}^{\mathrm{i}\theta} : |r - r_2| \le N^{-\alpha} \text{ and } |\theta| \le N^{-\alpha}\}$$

then for all large enough N, $u := r\mathrm{e}^{\mathrm{i}\theta} \in S_N$ implies $v := s\mathrm{e}^{\mathrm{i}\phi} := \{g(u)\}^{1/N} \in S_N$. Putting $c_1 := g(1)/2$ our assumptions imply that $c_1 > 0$. We have

$$\begin{aligned} |s^N - r_2^N| &= |\,|g(u)| - g(1)| \\ &\le |g(u) - g(1)|, \\ &\le c_2|u-1| \\ &\le c_2(|u - r_2| + |r_2 - 1|) \\ &\le c_3/N. \end{aligned}$$

Therefore

$$s^N \ge r_2^N - c_3/N = 2c_1 - c_3/N \ge c_1 > 0 \tag{27}$$

for all large enough N. Therefore

$$\sigma := \sum_{i+j=N-1} s^i r_2^j \ge N c_1^{(N-1)/N} \ge c_4 N.$$

Combining the above estimates yields

$$|s - r_2| \le \frac{c_3}{N\sigma} \le \frac{c_3}{c_4 N^2} \le N^{-\alpha}$$

for all large enough N.

We next observe that $v = s\mathrm{e}^{\mathrm{i}\phi}$ implies

$$\begin{aligned} |s^N \cos(N\phi) - g(1)| &= |\mathrm{Re}\,(v^N) - g(1)| \le |v^N - g(1)| \\ &= |g(u) - g(1)| \le c_2|u-1| \le c_5/N. \end{aligned}$$

Therefore

$$\cos(N\phi) > 0 \tag{28}$$

for all large enough N. Similarly

$$\begin{aligned} |s^N \sin(N\phi)| &= |\mathrm{Im}\,(v^N - g(1))| \le |v^N - g(1)| \\ &= |g(u) - g(1)| \le c_2|u-1| \le c_5/N. \end{aligned}$$

Using (27) we obtain

$$|\sin(N\phi)| \le \frac{c_5}{c_1 N}, \tag{29}$$

Combining (28) and (29) we deduce that

$$|\phi| \le \frac{2c_5}{c_1 N^2} \le N^{-\alpha}$$

for all large enough N. Therefore $v \in S_N$.

Having established that S_N is invariant under the map $u \to \{g(u)\}^{1/N}$ provided N is large enough, we now apply a contraction mapping argument within S_N. Let $z_j \in S_N$ and put $w_j := s_j \mathrm{e}^{\mathrm{i}\phi_j} := \{g(z_j)\}^{1/N}$ for $j = 1, 2$. Then

$$|w_1^N - w_2^N| = |g(z_1) - g(z_2)| \le c_6 |z_1 - z_2|.$$

Moreover

$$\begin{aligned}\sum_{i+j=N-1} w_1^i w_2^j &\ge \sum_{i+j=N-1} \mathrm{Re}\,(w_1^i w_2^j)\\ &= \sum_{i+j=N-1} s_1^i s_2^j \cos(i\phi_1 + j\phi_2)\\ &\ge N c_7\end{aligned}$$

where $c_7 > 0$. Therefore

$$|w_1 - w_2| \le c_6 |z_1 - z_2| / c_7 N \le |z_1 - z_2|/2$$

provided N is large enough. Since $u_2 \in S_N$, the contraction mapping principle implies that the sequence u_m defined by (25) converges as $m \to \infty$ to a solution $u \in S_N$ of $u^N = g(u)$, provided N is large enough.

The inclusion $u \in S_N$ implies $|u - r_2| \le 3N^{-\alpha}$. Putting $z := \mathrm{e}^{\mathrm{i}s_j} u$, we obtain $z^N = f(z)$ and $|z - a_j| \le 3N^{-\alpha}$ as required. $\square$

Note. Although we have proved that the eigenvalues of A all lie on or inside the unit circle asymptotically, this does not imply that $|\det(A)| \le 1$ asymptotically. Indeed $\det(A) = (-1)^{N-1} f(0)$ may be of any magnitude. If $|f(0)| > 1$ then the bound

$$|f(0)| = \left| \frac{1}{2\pi} \int_{-\pi}^{\pi} f(\mathrm{e}^{\mathrm{i}s})\, \mathrm{d}s \right| \le \frac{1}{2\pi} \int_{-\pi}^{\pi} \rho(s)\, \mathrm{d}s$$

implies that $\rho(s) > 1$ on average, so the eigenvalues close to the unit circle are actually slightly outside it, again on average.

Example 12. Theorem 11 is not applicable in the neighbourhood of any u such that $|u| = 1$ and $f(u) = 0$. However, the estimates in the theorem are local, so the conclusions are still valid for all the solutions of $z^N = f(z)$ that lie in any set of the form

$$S_{\alpha,\beta,\delta} = \{z : 1 - \delta \le |z| \le 1 + \delta \text{ and } \alpha \le \arg(z) \le \beta\},$$

provided α, β and $\delta > 0$ are chosen so that f does not vanish in $S_{\alpha,\beta,\delta}$. Fig. 2 shows the solutions of $z^N = 100(z - 1)$ when $N = 40$.

Acknowledgments

The first author acknowledges support under EPSRC grant number GR/R81756. The second author is very grateful to M. Zworski for pointing out this problem and its significance to her, and for his interest in and support of this work. She would also like to thank S. Nonnenmacher for his careful reading of a first draft. She has been supported by a postdoctoral grant of the École Polytechnique and by a CPAM grant at the University of California, Berkeley. Both authors would like to thank the referee for valuable comments.

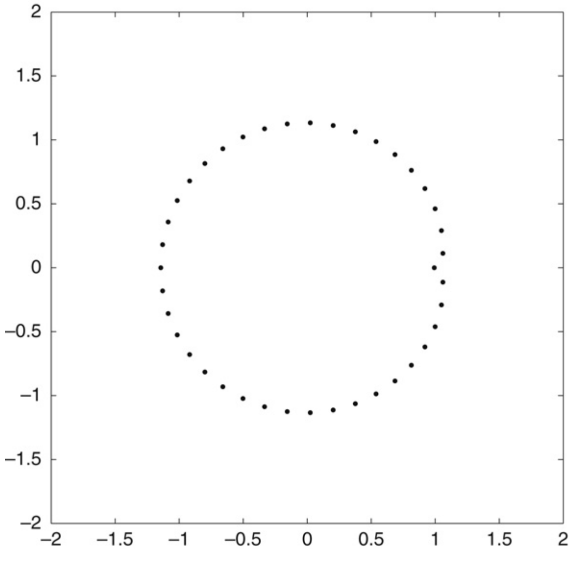

Fig. 2. Solutions of the polynomial equation of Example 12.

References

[1] N. Dencker, J. Sjöstrand, M. Zworski, Pseudospectra of semiclassical (pseudo-) differential operators, Comm. Pure Appl. Math. 57 (2004) 384–415.
[2] S.C. Reddy, P.J. Schmid, D.S. Henningson, Pseudospectra of the Orr-Sommerfeld operator, SIAM J. Appl. Math. 53 (1993) 15–45.
[3] J. Burke, M. Overton, Stable perturbations of nonsymmetric matrices, Linear Algebra Appl. 171 (1992) 249–273.
[4] M. Hager, J. Sjöstrand, Eigenvalue asymptotics for randomly perturbed non-selfadjoint operators, Math. Ann. 342 (1) (2008) 177–243.
[5] T. Kato, Perturbation Theory for Linear Operators, Springer-Verlag, New York, 1966.
[6] V.B. Lidskii, Perturbation theory of non-conjugate operators, USSR Comput. Math. Math. Phys. 6 (1966) 73–85.
[7] J. Moro, J.V. Burke, M.L. Overton, On the Lidskii-Vishik-Lyusternik perturbation theory for eigenvalues of matrices with arbitrary Jordan structure, SIAM J. Matrix Anal. Appl. 18 (1997) 793–817.
[8] Y. Ma, A. Edelman, Nongeneric eigenvalue perturbations of Jordan blocks, Linear Algebra Appl. 273 (1998) 45–63.
[9] L.N. Trefethen, M. Embree, Spectra and Pseudospectra: The Behaviour of Nonnormal Matrices and Operators, Princeton University Press, Princeton, NJ, 2005.
[10] E.B. Davies, Linear Operators and Their Spectra, Cambridge University Press, Cambridge, 2007.
[11] J. Sjöstrand, M. Zworski, Elementary linear algebra for advanced spectral problems, preprint 2003, http://www.math.berkeley.edu/~/ela.ps.gz.
[12] M. Hager, Instabilité spectrale semiclassique d'operateurs non-autoadjoints II, Ann. Henri Poincaré 7 (2006) 1035–1064.
[13] M.A. Beitia, I. Hoyos, I. Zaballa, The change of the Jordan structure under one row perturbations, Linear Algebra Appl. 401 (2005) 119–134.
[14] A. Edelman, Y. Ma, Non-generic perturbations of Jordan blocks, Linear Algebra Appl. 273 (1998) 45–63.
[15] R. Nevanlinna, Analytic Functions, Springer-Verlag, Berlin, New York, 1970.
[16] E.B. Davies, P.A. Incani, Spectral properties of matrices associated with some directed graphs, preprint, Feb. 2008.
[17] B. Simon, Rank one perturbations and the zeros of paraorthogonal polynomials on the unit circle, J. Math. Anal. Appl. 329 (2007) 376–382.